How to House a Mind Inside a Brain. Lessons from History.

Citation
Harrington, Anne. 2007. How to house a mind inside a brain. Lessons from history. EMBO reports 8(S1): S12–S15.

Published Version
doi:10.1038/sj.embor.7400974

EMBO reports

science & society

How to house a mind inside a brain

Lessons from history

Anne Harrington

The eighteenth-century French philosopher Voltaire once asked how it was that the great Newtonian heavens conform to the commands of physical law, but there remains in the universe "a little creature five feet tall, acting just as he pleases, solely according to his own caprice?" (Robinson, 1980).

The question was rhetorical, of course. A century earlier, Voltaire's compatriot René Descartes had famously offered humans an exemption from the natural order, by suggesting that causal principles applied to all intelligent behaviour in animals and all automatic behaviour in humans—such as snatching one's hand out of a flame—but that humans alone possessed a pure thinking substance, a conscious, wilful and rational soul created by God. This soul, Descartes said, directed all voluntary movements of the body, through the so-called animal spirits. Such thinking could no longer stand, Voltaire insisted. The time had come for humans to discover and acquiesce to their place in the natural scheme of things, regardless of the outcome.

Consequences aside, what would it mean, pragmatically, to put humans in their place in nature? From the beginning, the answer seemed clear: there must be no more exceptionalism. The human mind—the consciousness in each of us that peers through telescopes, scribbles calculations, falls in love, practices charity and ponders infinity—must be shown to be a product of the same impersonal forces that command the movements of the planets. This, in turn, meant that the new sciences must explain the functional relationship between human conscious experience and the small lump of living matter housed within the human skull, whose affinity for all things mental had been acknowledged even by Descartes: the human brain.

More than two centuries later, Voltaire's challenge still resonates. Today, we have high-tech brain-imaging machines, a general theory of the origin of life, a map of the human genome and a growing arsenal of pharmaceutical interventions to modulate or enhance our brain functions. However, in spite of all that, many of us still remain convinced that the truths of the brain—the logic and laws of this organ as a material entity—do not capture everything about humans. Any theory of our humanness must account for moral choice, existential passion and social contracts. It must be able to explain cathedrals, stock markets, wedding ceremonies, Shakespeare and people's willingness to die for their God. Moreover, most of us do not yet see how to relate the mind that seems to underlie those things to the material workings of brain processes. As the philosopher Colin McGinn put it, "The mind–body problem is the problem of understanding how [a] miracle is wrought" (McGinn, 1989).

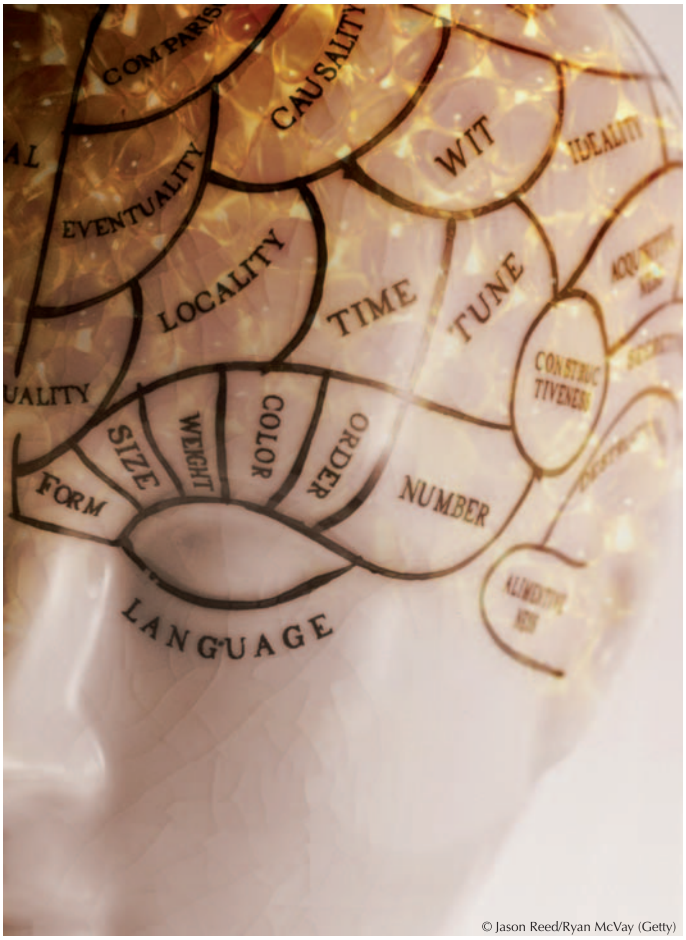
> **Any theory of our humanness must account for moral choice, existential passion and social contracts**

Is the problem of housing the mind inside the brain really so hard? Or is it possible that it feels so intractable because we are still trying to build on the clumsy or philosophically naive efforts of some of our forebears, whose labours lacked the finesse that we should rightly demand from today's brain sciences?

To find out, we should begin with the German physician Franz Joseph Gall at the start of the nineteenth century. Gall is well known for his organology or phrenology—that is, the attempt to determine character, personality traits or criminality from the skull's shape—but too often his system is regarded as a mere pseudoscience that is of no interest to today's students of the brain. That supposition is wrong: without Gall's interventions, the history of the brain sciences would almost certainly have developed differently. Gall offered two rules for how to house a mind inside a brain, which still dominate current thinking. The first rule was that to claim the human mind for science, you must first break it down into a fixed number of discrete functional units. The second rule was that having identified those functional units, you must then locate each of them in a discrete area of the brain.

These rules gambled on a bold assumption: that scientists and human brains had similar strategies for conceptualizing the functional building blocks of mental life. That is to say, brain localization theory under Gall—and beyond—began with conventional psychological categories such as 'language' and 'aggression', and hoped that these would represent discrete natural entities with a distinct physiological mechanism behind them. The gamble helped to launch powerful work in both the laboratory and the clinic, but was it rooted in a true premise? Even today, we cannot tell whether the British neurologist Sir Charles Scott Sherrington was right when he predicted that the scientific contributions of the localizers from Gall onwards would, when ultimately analysed, "resolve into components for which at present we have no names" (Sherrington, 1940).

When Gall tried to make a house for the mind in the brain and located each component of mental activity in its own room, he was challenging Christian dualistic theologies that required the soul to be both immaterial and undivided. His publications were placed on the Catholic list of prohibited books, and he was denied a Christian burial. Nevertheless, it is less appreciated that he never questioned the insight from the German philosopher Immanuel Kant that all living beings—in contrast to rocks and planets, for example—are animated by a sense of natural purpose (Zumbach, 1984). For Gall, the phrenological faculties—from language to maternal love—might indeed have been grounded in brain matter, but they also had distinctive roles in the larger economy of human life; each existed for a reason and did not simply result from a cause.



© Jason Reed/Ryan McVay (Getty)

By the late 1840s, however, this Kantian vision of multiple causalities was increasingly dismissed as the thinking of a generation of idealists. In particular, the so-called organic physicists in Germany—Hermann von Helmholtz, Emil Du Bois-Reymond, Ernst Brücke and Karl Ludwig—explicitly resolved that the time had come to build a science of life in which all explanations would be identical to the explanations in the physical sciences. As Du Bois-Reymond wrote to a friend at the time, "Brücke and I pledged a solemn oath to put into power this truth: no other forces than the common physical–chemical ones are active within the organism. In those cases which cannot at the time be explained by these forces one has either to find the specific way or form of their action by means of the physical–mathematical method, or to assume new forces equal in dignity to the chemical–physical forces inherent in matter, reducible to the forces of attraction and repulsion" (Sulloway, 1979).

When Gall tried to make a house for the mind in the brain […] he was challenging Christian dualistic theologies that required the soul to be both immaterial and undivided

From the synthesis of urea in the laboratory to the establishment of cell theory and mechanistic approaches to embryonic development, the nineteenth century saw one milestone discovery after another that seemed to support the biophysicists' cause. Of them all, however, none seemed to strike a firmer blow against exceptionalism than the first law of thermodynamics, which asserts that all forms of energy—mechanical, kinetic and thermal—are equivalent and can be transformed into one another. This implied that there was nothing 'extra' needed—or, indeed, allowed—to understand life, including the lives and minds of human beings. As the medical physiologist Rudolf Virchow described, "…the same kind of electrical process takes place in the nerve as in the telegraph line […] the living body generates its warmth through combustion just as warmth is generated in the oven; starch is transformed into sugar in the plant and animal just as it is in a factory" (Virchow, 1858).

What would it mean to begin with a vision of the brain as a 'parliament of localized faculties', as Gall had done, but then to go one step further and insist that the brain, the same as all living organs, generates mental life in the same way that an oven generates warmth? In 1874, the German psychiatrist Carl Wernicke—at the precocious age of 26—published a monograph on the problem of language loss (aphasia) and cerebral localization entitled *The Aphasic Symptom Complex* (Wernicke, 1874) that seemed to answer this question. It built on the work of the previous decade, and effectively amounted to a validation by mainstream medicine of at least a variant of Gall's approach but with a key difference.

Here is the background. In the early 1860s, the French anthropologist Paul Broca used clinical and anatomical evidence to persuade his colleagues, and eventually much of the larger international scientific community, that one of the phrenological mental faculties—what he called the "faculty of articulate language"—had a discrete seat in the brain, and that this lay in the third frontal convolution of the left frontal lobe of the human cortex (Broca, 1861, 1865). Wernicke, taking his starting point from the larger tradition that Broca had inaugurated, began with the observation that patients with damage to the left temporal lobe lose their capacity to comprehend language, even though they are generally able to speak fluently but do not make much sense. Wernicke contrasted this type of language disturbance with the syndrome identified by Broca in which people continue to understand what is said to them, but lose their capacity to express themselves.

Wernicke was not only completing Broca's work but also placing it in a new context. He proposed that one could not locate complex psychological units such as language in the brain, as Broca and his adherents—following Gall's lead—had indicated. All that could be localized were primitive 'memory traces' or impersonal records of sensory impressions and motor actions undertaken by the organism. On the basis of work on the basic neuroanatomy of spinal reflex action into the cortex, Wernicke proposed a model in which the rear of the cortex was specialized for processing and storing incoming sensory data, and the front part consisted of motor projections and centres that responded to sensory information with appropriate behaviours. Within the cortex itself, Wernicke claimed, all sensory–motor information mixed and communicated along 'fibres' that criss-crossed the cortex like telegraph lines. In fact, the brain was not a parliament of purposeful faculties; rather, it was a machine that generated complex mental processes directly from primitive non-thinking processes.

> **…the brain was not a parliament of purposeful faculties; rather, it was a machine that generated complex mental processes directly from primitive non-thinking processes**

Had Wernicke provided an explanation of how—to quote McGinn again—the "water of the physical brain is turned into the wine of consciousness"? At the beginning of the twentieth century, the Spanish neuroanatomist Santiago Ramon y Cajal expressed his doubts: "However excellent, every physiological doctrine of the brain based on localizations leaves us absolutely in the dark over the detailed mechanisms of the psychological acts" (y Cajal, 1960). In the end, the alchemy by which the brain became the mind would require knowledge of the "nature of the nervous wave, the energy transformations which [the mind] brings about or suffers at the moment when it is borne…".

Since the mid-eighteenth century, researchers had gathered evidence that this nervous energy was electrical in nature. Then, in the 1840s, du Bois-Reymond demonstrated nerve action potential that generated a constant current after stimulation. With the establishment—largely through y Cajal's efforts—of so-called neuronal theory (the idea that nerve cells are separated in space and communicate by some yet-to-be-determined process) scientists eventually developed an idea of a brain more dynamic than so far had been conceived. They began to visualize how electrical messages passing through the physical architecture of the nervous system might be purposefully directed, diverted, inhibited and augmented at different neuronal junctions, similar to a train having its direction and speed set and reset at various railroad switch points (y Cajal, 1954).

> **Even today, there are anomalies and internal divisions in the brave new world of brain science that indicate that we are not as close to fulfilling the hopes of the first bold generation of brain researchers as we might sometimes think**

In the early twentieth century, the English physiologist Sherrington recognized the empirical potential of neuronal theory. Working with dogs, he mapped the pathway taken by an impulse as it moved from a sensory receptor on the periphery—in this case, a tactile receptor on the skin—into the spinal cord and brain, and back out through a motor pathway where it produced a scratching response. The most important outcome of this work was a complex understanding of the reflex as a set of integrated interactions between electrical impulses on the one hand and chemical signals emitted at nerve junctions—which Sherrington named synapses—on the other. However, this did not mean that the brain worked like a train moving purposively along tracks with no driver and no switchmaster. In *Man on His Nature*, Sherrington painted a picture of the brain not as a mere mechanism, but rather as "an enchanted loom where millions of flashing shuttles weave a dissolving pattern, always a meaningful pattern though never an abiding one" (Sherrington, 1940). Moreover, he implied that the mind might be the weaver, because in spite of more than a century of efforts to house it firmly in the brain, the sciences still had to "regard the relation of mind to brain as not merely unresolved but still devoid of a basis for its very beginning".

These might have been the foolish words of a scientist past his prime and feeling the chill of his own mortality. Nevertheless, even if many would not have gone as far as Sherrington, some of the more thoughtful observers recognized that the expansionist ambitions of late nineteenth-century brain science had seen several setbacks. For instance, the French neurologist Jean-Martin Charcot had been humiliated when his efforts to make sense of hysteria as a neurological disorder according to anatomical principles had failed. In fact, the physical symptoms of hysteria—the paralyses and the loss of sensory function—turned out to be highly plastic and responsive to social cues. More specifically, one could make them disappear or change them using hypnotic suggestion.

None of that made sense in the world of brain functioning that Charcot knew, where brain disorders either were caused by stable lesions or were not. In the wake of the confusion, a nerve doctor from Vienna, Sigmund Freud, arrived on the scene and reinterpreted hysteria as a disease not of defective brain functioning but rather of a disordered biography and traumatic memories, thereby inaugurating a new way of—supposedly scientific—thinking about both the mind and its disorders, which declared the brain to be essentially irrelevant to the task.

Meanwhile, other discontented voices were being raised in neurology. Wernicke's radically mechanistic model of mind–brain functioning had come under fire—particularly in the German-speaking countries—by a group of dissatisfied scientists and clinicians who insisted that the mind was not something that could be broken down into primitive units and correlated to equally primitive physiological processes. The brain, they said, had properties as a whole that influenced lower-level functions in ways that neurology had little, if any, idea of how to conceive. Take, for example, the phenomenon of functional recovery after stroke or other brain damage. According to these dissatisfied researchers, the simple fact that brain-damaged people could improve over time, and regain speech and movement, was incompatible with the nineteenth-century model of the nervous system as a purely mechanical apparatus operating according to fixed laws of reflex and association. Machines do not repair themselves after suffering damage, and functions 'residing' in certain fixed regions of the brain cannot reappear if those regions are permanently destroyed. In this sense, the brain was not a machine—at least not in the way that people typically conceptualized the term.

To say this, however, was not to propose a return to the Cartesian view of the brain as an entity that stood in attendance on the needs and commands of a transcendent mind. It was simply to insist that any intellectually credible project to house a mind in a brain could not ignore anomalies that might demand some fundamental overhaul of core conceptual or methodological principles.

Today, this kind of radical talk perhaps seems quaint. We now believe that Freud is dead—or at least living out a declining reputation in academic departments of literature and cultural studies. Holistic approaches to the brain are out of fashion and most are as confident as ever of the basic 'divide and conquer' approach to putting the mind inside the brain.

Many of the reasons for this are more technological than conceptual, and are linked to the rise of neuroimaging machines and new psychopharmaceutical interventions. If it is true that specific drugs affect particular mental afflictions, does this not imply that Gall was right to see the mind as a mosaic of functional building blocks, even if we are now more sophisticated and appreciate the need to focus not only on neuroanatomy but also on neurochemistry? If different areas of the brain 'light up' on neuroimaging scans in real time when we think different thoughts, are we really so far from Gall's contemporary, Charles Bonnet, who proposed that anyone who thoroughly understood the functional anatomy of the brain would be able to read all the thoughts passing through it "as in a book" (Harrington, 1987).

Even today, there are anomalies and internal divisions in the brave new world of brain science that indicate that we are actually not as close to fulfilling the hopes of the first bold generation of brain researchers as we might sometimes think. In our time, ideas of hard-wired localization still exist with models of the nervous system as a self-updating system of dynamic neural nets or autopoietic processes. Research into the neurochemistry of the nervous system has raised questions as to whether the brain and nervous system even exist as an independent entity, or whether they represent a part of a system of interconnected physiological and biochemical processes, including those that regulate the gut and the immune system. In this sense, it seems possible that at least some of the processes that we consider an essential part of the mind are not housed strictly in the brain, but move freely across the body.

> **…it seems possible that at least some of the processes that we consider an essential part of the mind are not housed strictly in the brain, but move freely across the body**

None of this requires a retreat from the naturalism that Voltaire demanded of an enlightened age. Nevertheless, the fact that various projects in the brain sciences do not map seamlessly onto one another is important, as is the insistence with which many people—especially in the humanities and some of the social sciences—argue that our humanness needs to be understood not only as a product of the brain, but also as a product and expression of cultural, social and interpersonal forces. When social scientists and humanists resist the attempts of brain scientists to put the mind in its place, they are not just indulging in special pleading, expressing intellectual cowardice or trying to preserve space for something transcendent in us. Although these factors might sometimes be at play, history points to a more straightforward reason for the continuing discontent with, and internal divisions within, the brain sciences, even in these confident times. We are inheritors of a particular approach for housing a mind inside the brain—the 'divide and conquer' strategy—that has brought us far. However, history with all its dissension leads one to suspect that this strategy might not take us all the way home. The alchemy for transforming the water of the brain into the wine of consciousness might always elude us, but we can at least strive to fashion a bottle with sufficient room to encompass the fullness both of what we are and of how we experience ourselves to be.

REFERENCES

Broca P (1861) Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). *Bull Soc Anat (Paris)* **36:** 330–357

Broca P (1865) Du siège de la faculté du langage articulé. *Bull Soc Anthropol* **6:** 377–393

Harrington A (1987) *Medicine, Mind and the Double Brain*. Princeton, NJ, USA: Princeton University Press

McGinn C (1989) Can we solve the mind–body problem? *Mind* **98:** 349–366

Robinson DN (1980) *The Enlightened Machine*. New York, NY, USA: Columbia University Press

Sherrington CS (1940) *Man on his Nature*. Cambridge, UK: Cambridge University Press

Sulloway F (1979) *Freud: Biologist of the Mind*. Cambridge, MA, USA: Harvard University Press

Virchow R (1858) On the mechanistic interpretation of life. In *Disease, Life, and Man: Selected Essays*, Rather LJ (ed; 1958), pp 102–119. Stanford, CA, USA: Stanford University Press

Wernicke C (1874) Der aphasische Symptomenkomplex. Breslau, Germany: Crohn & Weigert

y Cajal SR (1954) *Neuron Theory or Reticular Theory? Objective Evidence of the Anatomical Unity of Nerve Cells*. Purkiss MU, Fox CA (transl). Madrid, Spain: Consejo Superior de Investigaciones Cientificas, Instituto Ramon y Cajal

y Cajal SR (1960) Anatomical and physiological considerations about the brain. In *Some Papers on the Cerebral Cortex*, von Bonin G (ed), p 275. Springfield, IL, USA: Thomas

Zumbach C (1984) *The Transcendent Science: Kant's Conception of Biological Methodology*. Boston, MA, USA: Kluwer Boston

Anne Harrington is Chair of the Department of the History of Science at Harvard University, Cambridge, MA, USA.
E-mail: aharring@fas.harvard.edu

doi:10.1038/sj.embor.7400974